What is claimed is:

1. An apparatus for processing three-dimensional video, comprising:

a storing means for storing video acquired with a predetermining video acquisition device;

a three-dimensional video generating means for converting a size and color of video transmitted from the storage;

an MPEG-4 control signal generating means for generating a Moving Picture Experts Group (MPEG)-4 object descriptor and a Binary Format for Scene (BIFS) descriptor;

an encoding means for encoding the three-dimensional video control signal and the MPEG-4 control signal inputted from the three-dimensional video generating means and the MPEG-4 control signal generating means, respectively through an MPEG-4 and encoding method, and outputting elementary stream (ES);

an MP4 file generating means for generating an MP4 file in conformity to an MPEG-4 system standards by receiving media data of the elementary stream outputted from the encoding means and the MPEG-4 control signal;

a packetizing means for extracting three-dimensional video media stream and the MPEG-4 control signal that are stored in the MP4 file generated in the MP4 file generating means, and generating and transmitting the extracted three-dimensional video media stream and the MPEG-4 control signal based on the MPEG-4 system standards;

a depacketizing means for receiving the packet stream transmitted from the packetizing means and depacketizing three-dimensional video data including a header and a payload;

a decoding means for decoding the data transmitted

from the depacketizing means and restoring three-dimensional video; and

a display means for displaying the video decoded in the decoding means.

2. The apparatus as recited in claim 1, wherein the three-dimensional video generating means acquires/generates three-dimensional video with the video acquisition device and the storing means, and converts the size and color of the acquired video.

3. The apparatus as recited in claim 1, wherein the MPEG-4 object descriptor includes information indicating whether the video inputted through the video acquisition device is binocular or multi-viewpoint three-dimensional video, information indicating the number of cameras/viewpoints of the inputted video, information indicating the number of media streams based on each viewpoint number, information indicating a two-dimensional/field shuttering/frame shuttering/polarized light display method with respect to binocular three-dimensional video, and information indicating a two-dimensional/panorama/stereoscopic display method with respect to multi-viewpoint three-dimensional video.

4. The apparatus as recited in claim 3, wherein the MPEG-4 control signal generating means generates the MPEG-4 object descriptor and the BIFS descriptor, and the MPEG-4 object descriptor including information on correlation between video and link structural information and includes information required for three-dimensional video while maintaining compatibility with a conventional object descriptor.

5. The apparatus as recited in claim 3, wherein

the decoding means decodes the three-dimensional video based on a system environment of a client and a display method selected by a user.

6. The apparatus as recited in claim 3, wherein the display means displays the decoded video and provides a user interface through rudimentary manipulation of the user to provide the user with the three-dimensional video.

7. A method for processing three-dimensional video in a video processing apparatus, comprising the steps of:

a) determining whether there is an access request from a client in a three-dimensional video transmitting server;

b) if there is no access request in the step a), maintaining a waiting mode or, if there is an access request, transmitting an initial object descriptor from the server to the client and establishing a session for a three-dimensional video service;

c) transmitting an MPEG-4 object descriptor or a Binary Format for Scene (BIFS) descriptor in the server upon receipt of a request for an object descriptor or a BIFS descriptor from the client; and

d) establishing a channel for transmitting three-dimensional video and transmitting the three-dimensional video upon receipt of a request for three-dimensional video from the client in the server, and restoring and displaying the three-dimensional video in the client.

8. The method as recited in claim 7, wherein the MPEG-4 object descriptor includes information indicating whether the video inputted through the video acquisition device is binocular or multi-viewpoint three-dimensional video, information indicating the number of cameras/viewpoints of the inputted video, information

indicating the number of media streams based on each viewpoint number, information indicating a two-dimensional/field shuttering/frame shuttering/polarized light display method with respect to binocular three-dimensional video, and information indicating a two-dimensional/panorama/stereoscopic display method with respect to multi-viewpoint three-dimensional video.

9. The method as recited in claim 7, wherein the information indicating whether the video inputted trough the predetermined video acquisition device is binocular/multi-viewpoint three-dimensional video occupies one bit and represents kind of three-dimensional video acquired according to the number and arrangement of cameras.

10. The method as recited in claim 9, wherein the information indicating the number of cameras/viewpoints of the inputted video, which occupies 10 bits, represents the number of viewpoints of the three-dimensional video and supports up to 1,024 viewpoints.

11. The method as recited in claim 10, wherein the information indicating the number of media stream according to each viewpoint number, which occupies one bit, represents the number of media stream according to each viewpoint number and presents a case where there are media elementary streams based on each viewpoint number and a case where the media elementary streams based on each viewpoint number are multiplexed and exist as one stream.

12. The method as recited in claim 10, wherein the information indicating a two-dimensional/field shuttering/frame shuttering/polarized light display method with respect to binocular three-dimensional video, which

occupies two bits, shows that the video inputted through the predetermined video acquisition apparatus is activated by the information indicating the binocular/multi-viewpoint three-dimensional video and represents a display method of the binocular three-dimensional video, which is a field shuttering display method, frame shuttering display method, polarized light display method, or two-dimensional display method.

13. The method as recited in claim 10, the information indicating a two-dimensional/panorama/stereoscopic display method with respect to multi-viewpoint three-dimensional video, which occupies two bits, is activated by the information indicating whether the video inputted through the predetermined video acquisition device is binocular/multi-viewpoint three-dimensional video and represents a multi-viewpoint three-dimensional video display method, which is a panorama display method, a two-dimensional display method, and a stereoscopic display method.